UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, STATION A, MONTRÉAL, QUÉBEC, H3C 3J7

ON THE CALIBRATION OF THE COMBINED DISTRIBUTION-ASSIGNMENT MODEL

by

S. Erlander, S. Nguyen, N.F. Stewart

Université de Montréal

Publication # 100 - Centre de recherche sur les transports

Juillet 1978

Dépôt légal - 4e trimestre 1978

ABSTRACT

In general the utilization of the combined distribution-assignment model requires that the deterrence factor be specified a priori. A natural approach to this question would be to obtain a value of the deterrence factor from an independent calibration of the gravity model. In this paper, we discuss the weakness of this approach and provide insights into the direct calibration of the combined distribution-assignment model.

1. Introduction

The traditional transportation planning process deals with four sequential stages: Trip generation, trip distribution, modal split, and assignment to the network. The process has to be iterated to find a stable solution since e.g. the assignment to the network influences the link costs and thus the travelling costs which in turn may cause a change in any of the three first stages. The combined model treats the second and the fourth stages simultaneously thus avoiding the sequential approach for these two stages.

The combined distribution-assignment model can be formulated in many ways (A short overview is given by Erlander, 1977). Here we shall adhere to the formulation used by Florian et al.(1975). In this case the objective function consists of a weighted sum of a certain (integral) function of the link costs and an entropy expression. The link cost function is introduced in order to produce equilibrium flows in the network whereas the entropy expression has the function of distributing the trips according to a gravity model. Thus this formulation may be looked upon as an entropy maximizing approach with equilibrium flows in the network. However, the same objective can be obtained if the entropy expression is first given as a constraint and then combined with the link cost function by means of a Lagrange multiplier. What follows will therefore be of more general interest and not restricted to the formulation explicitly treated here.

Before we proceed to give the mathematical formulation we shall briefly discuss the general background. We shall assume that every trip starts

from one out of a set of origin nodes, and finishes at one out of a set of destination nodes. The origin and destination nodes represent respectively the starting points and termination points for all trips within certain origin zones and destination zones. (The origin zones need not coincide with the destination zones.) Intrazonal trips and modal split are not treated in the formulation used here. It can be argued that intrazonal trips are treated in an unrealistic way; however, this is a separate question, which will be dealt with elsewhere. It has been shown that the model may be generalized to include modal split (Florian and Nguyen, 1978).

Consider the following formulation of the combined distribution assignment problem. Let N be a set of nodes, $A \subseteq N \times N$ a set of directed arcs and $G = (N,A)$ a graph representing a transportation network. Associated with each arc (i,j) A is a function c_{ij} which gives the unit cost of travel on the arc (i,j) as a function of the number of travellers on the arc; c_{ij} is assumed to be continuous and strictly monotonically increasing, and therefore integrable. There is also a set of origins $S \subseteq N$ and a set of destinations $T \subseteq N$.

The number of travellers leaving $s \in S$ is O_s, the number of travellers arriving at $t \in T$ is D_t, and $\sum_s O_s = \sum_t D_t$. All of these are assumed given and will remain fixed throughout:

$G = (N,A)$	$A \subseteq N \times N$	
$c_{ij}(\cdot)$	$(i,j) \in A$	
S,T	$S \subseteq N,\ T \subseteq N,$	
$O_s,\ D_t$	$O_s \geq 0,\ D_t \geq 0$	$s \in S,\ t \in T,\ \sum_s O_s = \sum_t D_t.$

We introduce the following notations related to numbers of travellers:

f_{ij}^s is the total number of travellers on arc (i,j) coming from s

$f_{ij} = \sum_s f_{ij}^s$ is the total number of travellers on arc (i,j)

g_{st} is the total number of travellers going from s to t.

We shall usually refer to the f's as flows and the g's as demands.

A descriptive model for the simultaneous prediction of the flows and the demands is

$$\text{Min } \alpha \sum_{(i,j)\in A} \int_0^{f_{ij}} c_{ij}(f)df + \sum_{s\neq t} g_{st} \log g_{st}$$

subject to the constraints

i) $\sum_{\substack{s \\ s\neq t}} g_{st} = D_t \quad t\in T$

ii) $\sum_{\substack{t \\ t\neq s}} g_{st} = O_s \quad s\in S$

iii) $$\sum_{k:(j,k)\in A} f_{jk}^s - \sum_{k:(k,j)\in A} f_{kj}^s + \begin{cases} -O_s & j=s \\ g_{sj} & j\in T \\ 0 & \text{oth.} \end{cases} = 0 \qquad \begin{matrix} s\in S \\ j\in N-\{s\} \ (j\in N\ j\neq s) \end{matrix}$$

iv) $f_{ij}^s \geq 0, \quad (i,j)\in A, \quad s\in S$

The parameter α must be specified in order to obtain a prediction for the f_{ij}^s, g_{st}. Once α has been specified, the prediction is unique, since the objective function is strictly convex. This will be made more precise in the following.

The term $\alpha \sum_{(i,j)\in A} \int_0^{f_{ij}} c_{ij}(f)df$ in the objective function is included along with the constraint iii) in order to obtain equilibrium flows; the term $\sum_{s\neq t} g_{st} \log g_{st}$ is included to obtain demands according to a gravity model.

In order to work easily with the model, it is useful to have an intuitive interpretation of what the model does. To begin with what the model does *not* do, we may observe that α does not quantify any trade-off between equilibrium solutions and maximum entropy solutions: the model always produces equilibrium solutions for the flows and gravity type solutions for the demands.

Rather, the model should be viewed as a locational analysis model which predicts (say) where people will choose to live and where people will choose to work, and the corresponding travel demands g_{st}. The parameter α quantifies the extent to which people take into account the (equilibrium) cost of travel when making decisions. When $\alpha \to 0$ they take no account of the cost of travel, and the demands g_{st} will be the uniform demands produced by entropy maximization with no constraint on the total cost; the flow f_{ij}^s will be the equilibrium flows associated with these demands. On the other hand, when $\alpha \to \infty$ the model produces demands which correspond to a locational choice which minimizes the (equilibrium) cost of travel; again the flows f_{ij}^s will be the equilibrium flows associated with these demands.

The Kuhn-Tucker conditions associated with the above minimization problem are:

$$\log g_{st} + \alpha u_{st} + \beta_s + \gamma_s + 1 = 0 \qquad s \neq t$$

(obtained by differentiating w.r.t. g_{st})

and

$$\left.\begin{aligned} & u_{si} + c_{ij}(f_{ij}) \geq u_{sj} \\ & u_{si} + c_{ij}(f_{ij}) = u_{sj} \quad \text{if} \quad f_{ij}^s > 0 \end{aligned}\right\} \quad \begin{matrix} s \in S \\ (i,j) \in A \end{matrix}$$

(obtained by differentiating w.r.t. f_{ij}^s).

The multiplier β_s corresponds to constraint ii), γ_t to constraint i), and u_{sj} to constraint iii). The multiplier u_{sj} may be interpreted as the cost at equilibrium of the least expensive path of travel from s to j. Since the objective function is (strictly) convex and the constraints are linear the Kuhn-Tucker necessary conditions are also sufficient. The Kuhn-Tucker constraint qualification is satisfied because of the linearity of the constraints.

In the following we shall use the notation C D A to denote the combined distribution-assignment model just formulated.

2. Calibration of the model

The C D A model as given in the preceeding section contains one free parameter, namely the parameter α, which has to be specified; the model has to be calibrated.

When confronted with a model containing a free parameter, engineers have traditionally eliminated it, or tried to eliminate it, by insisting that the model reproduce supplementary information when the model is applied in an earlier year for which such supplementary information is available. For

example, in the case of the classical gravity model it is assumed that $\hat{u}_{st}$ and $\hat{C}$ are available in an earlier year, where $\hat{C}$ is the total cost of travel, and it turns out that for values of $\hat{C}$ between certain well defined limits this assumption is sufficient to specify the free parameter uniquely (Evans, 1971). Can we do something similar with the C D A model?

A natural first approach to the calibration of the C D A model is to assume available $\hat{u}_{st}$ and $\hat{C}$, or at least estimates of these quantities, and to calibrate the classical gravity model to obtain a value of α:

$$\underset{0 \le g_{st}}{\text{Min}} \quad \sum_s \sum_t g_{st} \log g_{st}$$

subject to the constraints i), ii) and $\sum_s \sum_t \hat{u}_{st} g_{st} = \hat{C}$

Now, if we find the equilibrium flow corresponding to the resulting demands g_{st}, we shall find in general that the resulting equilibrium travel costs differ from the $\hat{u}_{st}$ which were used in the calibration of the gravity model. Consequently, if we take the value of α obtained above and solve the corresponding version of the C D A model, it will in general not reproduce the $\hat{C}$ used in the calibration of the gravity model.

This seems to be an unsatisfactory situation, since one of the principal reasons for introducing the C D A model was to avoid the difficulty concerning the u_{st} which occurs when distribution and assignment are done sequentially. The model avoids the difficulty in the prediction year, but not in the earlier year (or "base year") when the model is calibrated. It is natural therefore to ask whether the C D A model can be calibrated directly. The answer is that it depends upon on how the calibration is

done. It is not possible to calibrate on total costs as in the classical gravity model case. The reason for this is that total costs are not uniquely determined by the parameter α. This will be shown by giving counterexamples further on. On the other hand it is possible, as will be shown in the next section, to calibrate on

$$F \equiv \sum_{(i,j)} \int_0^{f_{ij}} c_{ij}(f)df, \quad \text{or}$$

$$H \equiv \sum_{s \neq t} g_{st} \log g_{st},$$

since both F and H are strictly monotone functions of α except possibly for certain intervals of the positive real axis where they are constant and where the flows and demands are constant as well.

3. Monotonicity of the F and H functions

We shall now study the monotonicity of the functions

$$F(\alpha) = \sum_{(i,j) \in A} \int_0^{f_{ij}(\alpha)} c_{ij}(f)df, \quad \text{and}$$

$$H(\alpha) = \sum_{s \neq t} g_{st}(\alpha) \log g_{st}(\alpha),$$

where $f_{ij}(\alpha)$ and $g_{st}(\alpha)$ are the optimal solutions to the minimization problem

$$\text{Min } \alpha \sum_{(i,j) \in A} \int_0^{f_{ij}} c_{ij}(f)df + \sum_{s \neq t} g_{st} \log g_{st}$$

subject to the constraints i),...,iv).

The optimal flows $f_{ij}(\alpha)$ and the optimal demands $g_{st}(\alpha)$ are well defined functions of α for $0 < \alpha < \infty$. If $\alpha = 0$ the flows may not be unique, and a more elaborate definition is required. In this paper we restrict our attention to $0 < \alpha < \infty$.

Lemma 1

The minimization problem

$$\text{Min } \alpha \sum_{(i,j)\in A} \int_0^{f_{ij}} c_{ij}(f)df + \sum_{s\neq t} g_{st} \log g_{st}$$

subject to the constraints i),...,iv)

has a unique optimal solution $f_{ij}(\alpha)$, $g_{st}(\alpha)$

Proof

The assertion follows from the strict convexity of the objective function in f_{ij} and g_{st} and the convexity of the feasible region.

We shall now prove that F and H are monotone but not necessarily strictly monotone.

Lemma 2

For $0 < \alpha < \infty$, $F(\alpha)$ is monotonically decreasing, and $H(\alpha)$ is monotonically increasing.

Proof

For given positive constants α_1 and α_2, we have

$$\alpha_1 F(\alpha_1) + H(\alpha_1) \leq \alpha_1 F(\alpha_2) + H(\alpha_2), \quad \text{and}$$

$$\alpha_2 F(\alpha_2) + H(\alpha_2) \leq \alpha_2 F(\alpha_1) + H(\alpha_1).$$

Therefore,

$$(1) \quad \alpha_1[F(\alpha_2) - F(\alpha_1)] \geq H(\alpha_1) - H(\alpha_2) \geq \alpha_2[F(\alpha_2) - F(\alpha_1)]$$

and

$$(\alpha_2 - \alpha_1)[F(\alpha_2) - F(\alpha_1)] \leq 0.$$

It follows that

$\alpha_2 > \alpha_1 \Rightarrow F(\alpha_2) - F(\alpha_1) \leq 0$

and

$\alpha_2 > \alpha_1 \Rightarrow H(\alpha_2) - H(\alpha_1) \geq 0$

Lemma 3

If either F or H is constant on an interval $[\alpha_1,\alpha_2]$, $0 < \alpha_1 < \alpha_2$, then all of the following are constant for $\alpha \in [\alpha_1,\alpha_2]$:

$F(\cdot)$, $H(\cdot)$,

$f_{ij}(\cdot)$ all $(i,j) \in A$,

$g_{st}(\cdot)$ all $s \neq t$

Proof. Suppose first that $F(\alpha_1) = F(\alpha_2)$. It then follows from formula (1) that $H(\alpha_1) - H(\alpha_2) \geq 0$. But according to Lemma 2 H is monotonically increasing. Hence $H(\alpha_1) = H(\alpha_2)$. Since $\alpha_1 > 0$ it follows immediately from (1) that $F(\alpha_2) - F(\alpha_1) \geq 0$, and hence from the monotonicity, $F(\alpha_1) = F(\alpha_2)$. Now both F and H are strictly convex functions of $f_{ij}(\alpha)$ and $g_{st}(\alpha)$, respectively. If neither function changes on $[\alpha_1,\alpha_2]$, it must be that none of the $f_{ij}(\alpha)$, $(i,j) \in A$, nor the $g_{st}(\alpha)$, $s \neq t$ change either.

We are now prepared to give our main result, i.e. that it is possible to calibrate the C D A model on F or H.

Theorem. Let $\hat{F}$ or $\hat{H}$ be given, and let α be a solution to one of the equations

$$F(\alpha) = \hat{F} \quad \text{or} \quad H(\alpha) = \hat{H},$$

where $F(\alpha)$ and $H(\alpha)$ denote the values of F and H obtained at the optimal solution to the problem

$$\text{Min} \quad \alpha \sum_{(i,j)\in A} \int_0^{f_{ij}} c_{ij}(f)df + \sum_{s\neq t} g_{st} \log g_{st}$$

subject to (i),...,(iv).

Then the corresponding flows $f_{ij}(\alpha)$ and demands $g_{st}(\alpha)$ are unique.

Proof: The assertion follows from the Lemmas.

The functions F and H are thus monotone, but perhaps not strictly monotone. On any interval where one or the other is constant, all of the demands and flows are constant. Thus we either get a unique α for a given value of F or H, or else (if it is not unique) the underlying problem is so degenerate that the choice of α is irrelevant - any of the choices of α which give the given value of F or H will produce the same demands and flows.

4. Non-uniqueness of α given total cost

Given an estimate of the total cost $\hat{C}$ in the base year, is there a unique value of α which, when the corresponding version of the C D A model is solved, reproduces $\hat{C}$? The answer is "no". We shall show, by giving a counterexample, that the total cost function

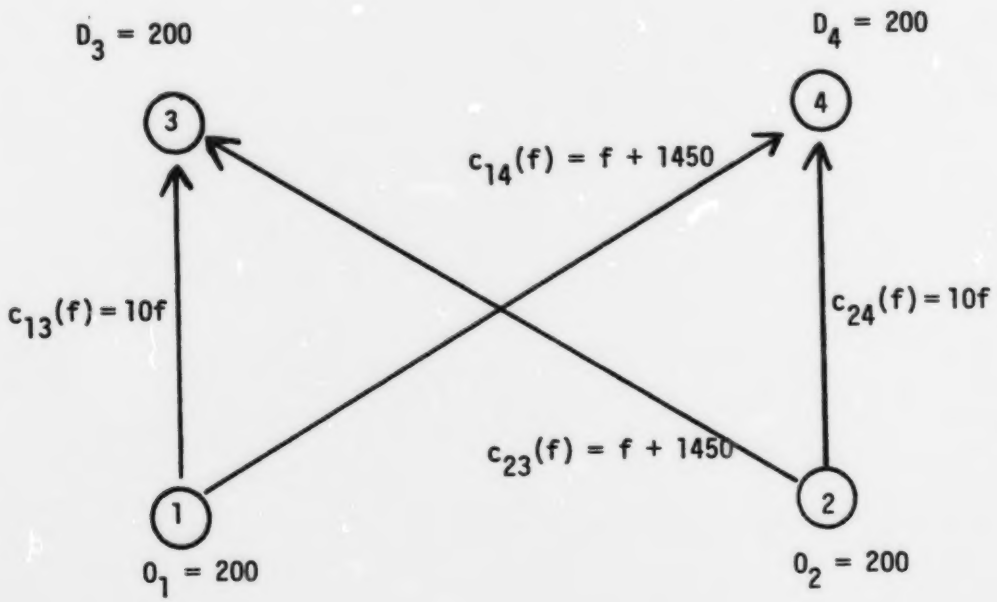
$$C(\alpha) \equiv \sum_{s \neq t} u_{st}(\alpha)\, g_{st}(\alpha) \equiv$$

$$\equiv \sum_{(i,j) \in A} c_{ij}(f_{ij}(\alpha))\, f_{ij}(\alpha),$$

where the $u_{st}(\alpha)$, $g_{st}(\alpha)$ and the $f_{ij}(\alpha)$ are the values produced by the C D A model, is not always monotone. This means that the C D A model can in general not be uniquely calibrated on total cost.

We shall first construct an example where $C(\alpha)$ increases, then one with decreasing $C(\alpha)$, and finally by combining the two obtain an example where $C(\alpha)$ first decreases and then increases. In this combined example the cost functions $c_{ij}(\cdot)$ are convex. To show that convexity is not necessary we shall give one example with concave cost functions $c_{ij}(\cdot)$. In our examples two values of α give the same value to the total cost function. It will be clear however that more complicated (although conceptually not more difficult) examples could be constructed with many values of α giving the same value to $C(\cdot)$. This means clearly that the C D A model can not be uniquely calibrated on total cost.

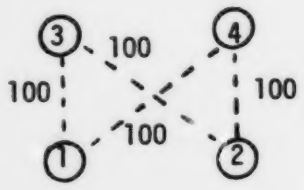
The key to understanding the difficulty with the proposed direct calibration on total cost is to be found in the following example; the example is just an application of the well known fact that individuals acting solely for their own benefit may produce a solution which is system sub-optimal.



When $\alpha = 0^+$ the model produces the following limiting values for the demands:

g_{st}	3	4
1	100	100
2	100	100

denoted



which comes from the minimization of $\sum_{s \neq t} g_{st} \log g_{st}$

subject to $\sum_t g_{st} = 200 \quad s = 1,2$

$\sum_s g_{st} = 200 \quad t = 3,4$

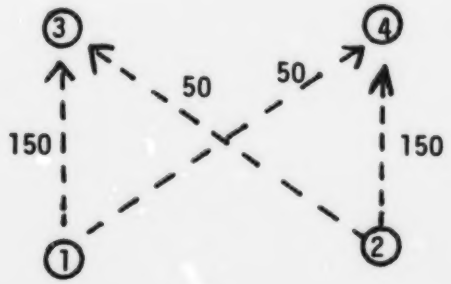
The equilibrium assignment, given these demands is

$$f_{ij} = g_{ij} \qquad i=1,2 \qquad j=3,4$$

(In fact, the demands specify the flows in this way, independently of the costs, because of the special nature of the network.) The total cost of travel is

$$C(0^+) = 100(1000) + 100(1550) + 100(1000) + 100(1550) = 510,000.$$

When $\alpha > 0$, people living at ① will pay more attention to the (equilibrium) cost of travel, and more of them will choose to work at ③ rather than at ④, because the cost of travel is lower. Indeed when $\alpha \to \infty$, the model produces demands (and therefore flows) with the limiting values:



for which the total cost of travel is

$$\lim_{\alpha \to \infty} C(\alpha) = 150(1500) + 50(1500) + 150(1500) + 50(1500) = 600,000.$$

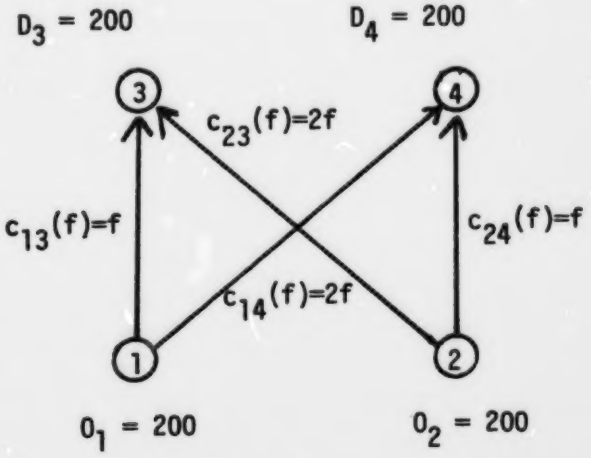
Although the cost of travel on (1,3) and (2,4) was lower, the marginal cost was higher, and those who shifted their place of work (or their residence)

have made travel much more expensive for their neighbours.

For $\alpha = \frac{\log(5/3)}{275}$, an elementary argument shows that $C(\alpha) = 541{,}250$. (There is nothing special about this value of α except that it is the value which produces the demands

and this will be useful later).

We turn now to the second example, which is intuitively the most appealing, given the results for the classical gravity model. Here it is sufficient to take the same network as that just used, but give the smaller cost function $c_{13}(\cdot)$ and $c_{24}(\cdot)$ a lower marginal cost too. For example



For $\alpha = 0^+$ we obtain uniform limiting demands, and

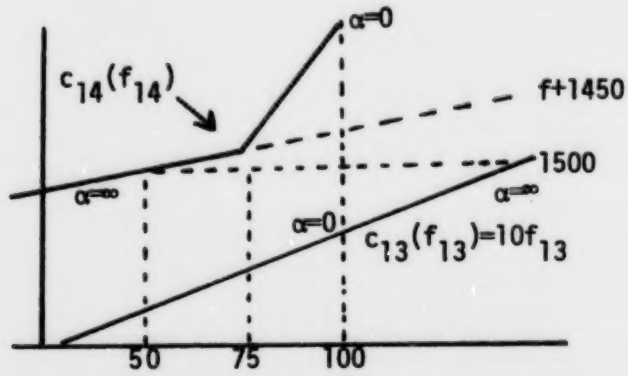
$$C(0^+) = 100(100) + 100(200) + 100(100) + 100(200) = 60{,}000$$

When $\alpha \to \infty$ we obtain the limiting values $g_{13} = f_{13} = \frac{400}{3}$, and

$$\lim_{\alpha \to \infty} C(\alpha) = \frac{400}{3}\left(\frac{400}{3}\right) + \frac{200}{3}\left(\frac{400}{3}\right) + \frac{400}{3}\left(\frac{400}{3}\right) + \frac{200}{3}\left(\frac{400}{3}\right) = 53{,}333\tfrac{1}{3}$$

This time the shifting benefits everyone.

We now combine the structure of the two examples in such a way that the shifting from (1,4) to (1,3) (and from (2,3) to (2,4)) is at first beneficial for everyone, then antisocial. We take c_{14} and c_{23} to be the same, and similarly for c_{13} and c_{24}:



Specifically:

$$c_{13}(f) = c_{24}(f) = 10f,$$

$$c_{14}(f) = c_{23}(f) = \begin{cases} f + 1450 & \text{for } 0 \le f \le 75, \\ 20f + 25 & \text{for } 75 \le f. \end{cases}$$

The minimization problem for $\alpha = 0^+$ gives the symmetric demands of 100 for each O-D pair: the total cost is:

$$C(0^+) = 100(2025) + 100(1000) + 100(2025) + 100(1000) = 605{,}000$$

and at $\alpha = 0^+$ we have $c_{14}(f_{14}) = 2025$

$$c_{13}(f_{13}) = 1000.$$

As α increases from zero, the demand shifts until $f_{14} = 75$, $f_{23} = 75$,

$$\alpha = \frac{\log\left(\frac{5}{3}\right)}{275}, \text{ and the total cost is}$$

$$C\left(\frac{\log\left(\frac{5}{3}\right)}{275}\right) = 541{,}250.$$

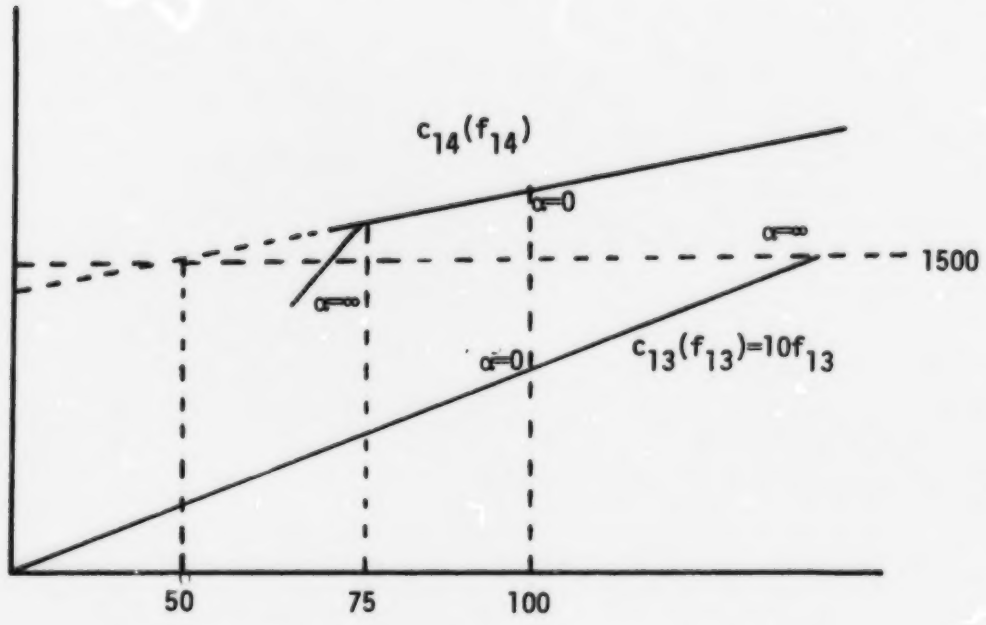
The shifting so far has been from (1,4) to (1,3) and it has reduced the total cost. However, if we increase α further, we are back to the first example: the shifting is still from (1,4) to (1,3), but the total cost will begin to increase until

$$C(\infty) = 150(1500) + 50(1500) + 150(1500) + 50(1500)$$
$$= 600{,}000.$$

Consequently we may have two different values of α with completely different flows, for which the C D A model produces the same value of C.

In the formulation of the model we assumed only that the c_{ij} were continuous and strictly monotone. However, it is clear that the discontinuity in the derivative of c_{14} played no crucial role in the example, and we could give such an example even if c_{14} were required to be very smooth.

To obtain an example with concave cost functions it is clearly sufficient to modify c_{14} (and c_{23}) in the following way:



Until $f_{14} = 75$ the shifting will be antisocial; after that, it will be beneficial.

5. Conclusion

Utilization of the CDA model requires that the deterrence factor α be determined *a priori*. A natural approach to this question would be to obtain a value of α from an independant calibration of the gravity model, using one of the several existing techniques. However, this approach is difficult to justify, since one of the principal reasons for introducing the CDA model was to avoid the inconsistency between the interzonal travel times used for the calibration of the distribution model, and the travel times resulting from a subsequent equilibrium assignment of the network.

The authors believe therefore that the deterrence factor α should be

obtained directly from the CDA model. The present paper shows which approaches to direct calibration are feasible and which are not, by showing that the function F (the sum of the integrals of the link travel time functions), and the negative entropy function H, are monotonic functions of α, but that the total cost function C is not in general monotonic. These results show that the CDA model may be calibrated directly, given either an estimate of F or an estimate of H. The process consists of determining a consistent value of α (using, for example, a binary search scheme), such that the model's value of F (or H) agrees, to within a given tolerance, with the estimated value in the base year.

The estimation of F or H given base year data does not raise any theoretical problem, but the practical statistical problems associated with such estimation remain to be studied, including both the question of sample size and the question of statistical independance of samples. In the case of the function F an alternative approach appears very promising: if we assume available the interzonal travel times, and compute the unique equilibrium flow associated with these travel times, we may then compute F. (This latter approach is similar to a method proposed for a different purpose by Nguyen (1977).) The authors intend to compare these approaches to the estimation of F and H (and hence α) when they are applied using real data.

REFERENCES

Erlander, S. (1977) Accessibility, entropy and the distribution and assignment of traffic. Transportation Research 2, 149-153.

Evans, A.W. (1971) The calibration of trip distribution models with exponential or similar cost functions. Transportation Research 5, 15-38.

Florian, M., Nguyen, S. and Ferland, J. (1975) On the combined distribution-assignment of traffic. Transportation Science 9, 43-53.

Florian, M. and Nguyen, S. (1978) A combined trip distribution, modal split and trip assignment model. To appear in Transportation Research.

Nguyen, S. (1977) Estimating an O-D matrix from network data: a network equilibrium approach. Publication # 60, Centre de recherche sur les transports, Université de Montréal.